MODELING LOAD VARIATIONS WITH ARTIFICIAL NEURAL NETWORKS TO IMPROVE ON-WAFER OSLT CALIBRATIONS*

Jeffrey A. Jargon[1], Pete Kirby[2], K. C. Gupta[3], Lawrence Dunleavy[2], and Tom Weller[2]

[1]National Institute of Standards and Technology
RF Electronics Group, 325 Broadway, Boulder, CO 80303
e-mail: jargon@boulder.nist.gov

[2]Wireless and Microwave (WAMI) Program
University of South Florida, Tampa, FL 33620

[3]Center for Advanced Manufacturing and Packaging of
Microwave, Optical, and Digital Electronics (CAMPmode)
University of Colorado at Boulder, Boulder, CO 80309

ABSTRACT

We demonstrate that on-wafer open-short-load-thru (OSLT) calibrations of vector network analyzers can be improved by applying artificial neural networks (ANNs) to model the correlation between DC resistance and RF variations in load terminations. The ANNs are trained with measurement data obtained from a benchmark multiline thru-reflect-line (TRL) calibration. The open, short, and thru standards do not vary significantly from wafer to wafer, so we also model these standards using ANNs trained with calibrated measurement data chosen from an arbitrary wafer. We assess the accuracy of five OSLT calibrations with varying load terminations using the ANN-modeled standards, and find that they compare favorably (a difference of less than 0.04 in magnitude at most frequencies) to the benchmark multiline TRL calibration over a 66 GHz bandwidth. We demonstrate that ANN models offer a number of advantages over using calibrated measurement files or equivalent circuit models, including ease of use, reduced calibration times, and compactness.

I. INTRODUCTION

Multiline thru-reflect-line (TRL) is a highly accurate means of VNA calibration and is especially useful for on-wafer environments, since the characteristic impedance can be calculated from dimensional measurements of the standards, which simply consist of a number of transmission lines of varying line lengths and a highly reflective termination [1]. The disadvantages of this calibration method are that it requires a lot of real estate on the wafer, due to the numerous long lines required for an accurate calibration, and the different lengths of lines necessitates changing the separation between probes during the calibration process. Consequently, compact calibration kits, such as open-short-load-thru (OSLT) [2], are usually preferred for on-wafer applications. The trade-off is that the kits with smaller, lumped-element artifacts tend to be less accurate, since it is more difficult to calculate the reflection coefficients of the standards. But, if the compact calibration kits can be characterized using a benchmark calibration, such as multiline TRL, it is possible to perform an accurate on-wafer OSLT calibration.

Once the OSLT standards on a given wafer are characterized, we must decide whether to develop a model for each of the standards or to directly use the measurement data obtained from the benchmark calibration. Recently, Jargon et al. [3] applied artificial neural networks (ANNS) to improve the modeling of on-wafer OSLT standards. They showed that ANN models offer a number of advantages over the use of calibrated measurement data files or equivalent circuit models, namely, the following: (1) they do not require detailed physical models, (2) calibration times can be reduced since only a few training points are required to accurately model the standards, (3) ANN model descriptions are much more compact than large measurement files, (4) ANN models, trained on only a few measurement points, can be much more accurate than direct calibrations, when limited data are available, and (5) they are less susceptible to the noise inherent in measured data. The assumption made in this work was that the standards can be reproduced from wafer to wafer with little variation.

Kirby et al. [4] studied variations in OSLT standards from wafer to wafer on a CPW calibration set designed for GaAs substrates, and found that open, short, and thru standards can be reproduced with minimal variance, but that load standards exhibit a significance difference among the wafers they studied. Furthermore, they discovered that RF variations in the load terminations correlate directly to their measured DC resistances.

We have expanded upon the method of [4] by implementing ANNs to model RF variations in the load standards as a function of DC resistance. The following sections describe our implementation of ANNs to model the on-wafer OSLT calibration standards, and our assessment of the accuracy of five OSLT calibrations with varying load terminations using the ANN-modeled standards.

II. ARTIFICIAL NEURAL NETWORKS

ANNs have been applied to diverse areas such as speech and pattern recognition, financial and economic forecasting, telecommunications, and nuclear power plant diagnosis, and have just recently been introduced into the area of microwave engineering [5-8]. In particular, researchers have successfully used ANNs to model microstrip vias [9], packaging and interconnects [10], spiral inductors [11], MESFET devices [12], CPW circuit components [13], effective dielectric constant of microstrip lines [14], and HBT amplifiers [15], to name just a few.

The ANN architecture used in this work is a feed-forward, three-layer perceptron structure (MLP3) consisting of an input layer, a hidden layer, and an output layer, as shown in Figure 1. The hidden layer allows for complex modeling of input-output relationships. The mapping of these relationships is given by [9]

$$\mathbf{Y} = g[\mathbf{W}_2 \cdot g(\mathbf{W}_1 \cdot \mathbf{X})],$$

where **X** is the input vector, **Y** is the output vector, and $\mathbf{W}_1$ and $\mathbf{W}_2$ are the weight matrices between the input and hidden layers and between the hidden and output layers, respectively. The function $g(u)$ is a nonlinear sigmoidal activation function given by

$$g(u) = \frac{1}{1 + \exp(-u)},$$

where u is the input to a hidden neuron. An MLP3, with one hidden sigmoidal layer, is able to model almost any physical function accurately, provided that a sufficient number of hidden neurons are available [8].

ANNs learn relationships among sets of input-output data that are characteristic of the device or system under consideration. After the input vectors are presented to the input neurons and output vectors are computed, the ANN outputs are compared to the desired outputs and errors are calculated. Error derivatives are then calculated and summed for each weight until all of the training sets have been presented to the network. The error derivatives are used to update the weights for the neurons, and training continues until the errors reach prescribed values. In this study, we utilized software developed by Zhang et al. [16] to construct our ANN models.

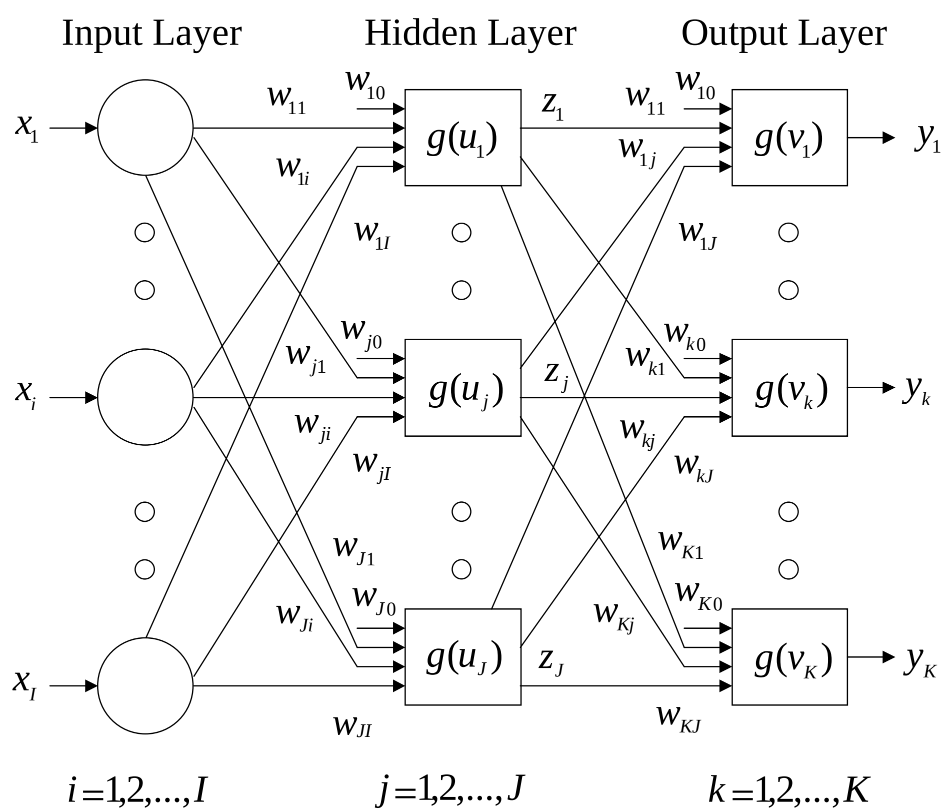

Figure 1. Artificial neural network architecture.

III. MODELING THE STANDARDS

In this study, the OSLT and multiline TRL standards and devices were constructed of CPW transmission lines fabricated from 4.5 μm plated gold on a 625 μm thick GaAs. The load terminations were composed of TiWN (titanium tungsten nitride) thin film resistive material [4]. The four line standards included a thru line and three additional lines that were 0.9552, 1.239, and 1.764 mm longer. All of the standards were measured using on-wafer probes. For each standard, we measured scattering parameters at 165 frequencies from 1 to 67 GHz.

Since the open, short, and thru standards did not vary significantly from wafer to wafer we modeled these standards with ANNs using calibrated measurement data chosen from an arbitrary wafer. The ANN architecture for the open, short, and thru standards consisted of one input (frequency) and two outputs (the real and imaginary components) for each measured scattering parameter. Since we measured reflection coefficients for the two terminations at both ports and all four scattering parameters of the thru connection, we ended up with eight ANN models, excluding the load. From our previous study in [3], we determined that 5 neurons were sufficient for the hidden layer. We trained each model of the standards using all 165 frequencies since we already had the data on hand. Figures 2-3 show the magnitude and phase of S_{11} of both measured and ANN model data for the open and short standards, respectively. Figure 4 shows the magnitude and phase of S_{11} and S_{21} using the measured and ANN model data for the thru standard. Notice that the ANN models for each standard follow the trends of the measured data, but avoid the scatter of the multiline TRL calibrated measurements. Whether or not this scatter is real, we see that ANNs follow general trends, but omit the noise, which is usually desirable in a model.

The ANN architecture for the load standards consisted of two inputs (frequency and DC resistance) and two outputs (the real and imaginary components) for the impedance parameters at each port. We were unable to generate one model that included both ports due to a systematic difference between the load measurements at port 1 and port 2, so we settled on separate models for each port. Ten neurons were chosen for the hidden layers since the ANN models for the loads included an additional input compared to the other standards. The measured DC resistances for the loads are listed in Table 1. For each port, we trained the models using 3 of the 5 loads. We chose loads 1, 4, and 5 since load 1 had the lowest DC resistance, load 5 had the highest, and load 4 had an intermediate value. It is important to train ANNs at the expected boundary values of the input parameter space in order to ensure good performance of the model [6]. By purposely not training the ANN with loads 2 and 3, we could test how effective the model behaved at other DC resistances. Figure 5 shows the real and imaginary components of Z_{11} of both measured and ANN model data for the 5 load standards. Likewise, Figure 6 shows the real and imaginary components of Z_{22} of both measured and ANN model data for the 5 load standards. Once again, we see that ANNs follow general trends while omitting the noise.

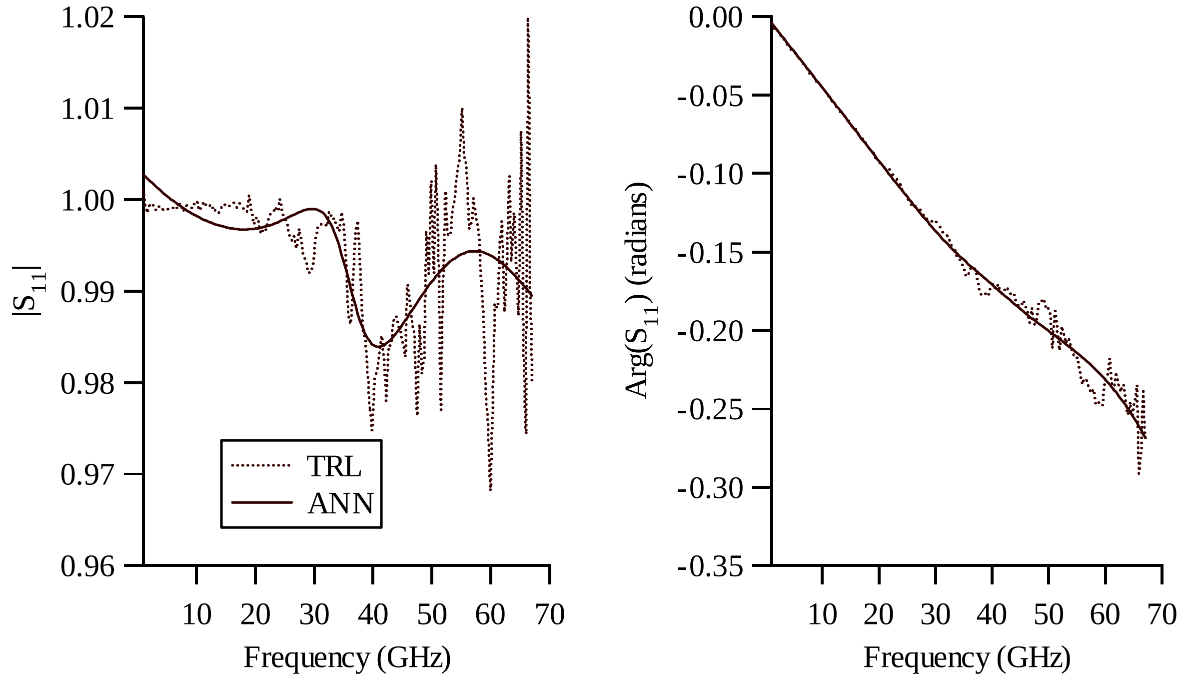

Figure 2. Magnitude and phase of S_{11} for the open standard measured by multiline TRL and modeled by an ANN.

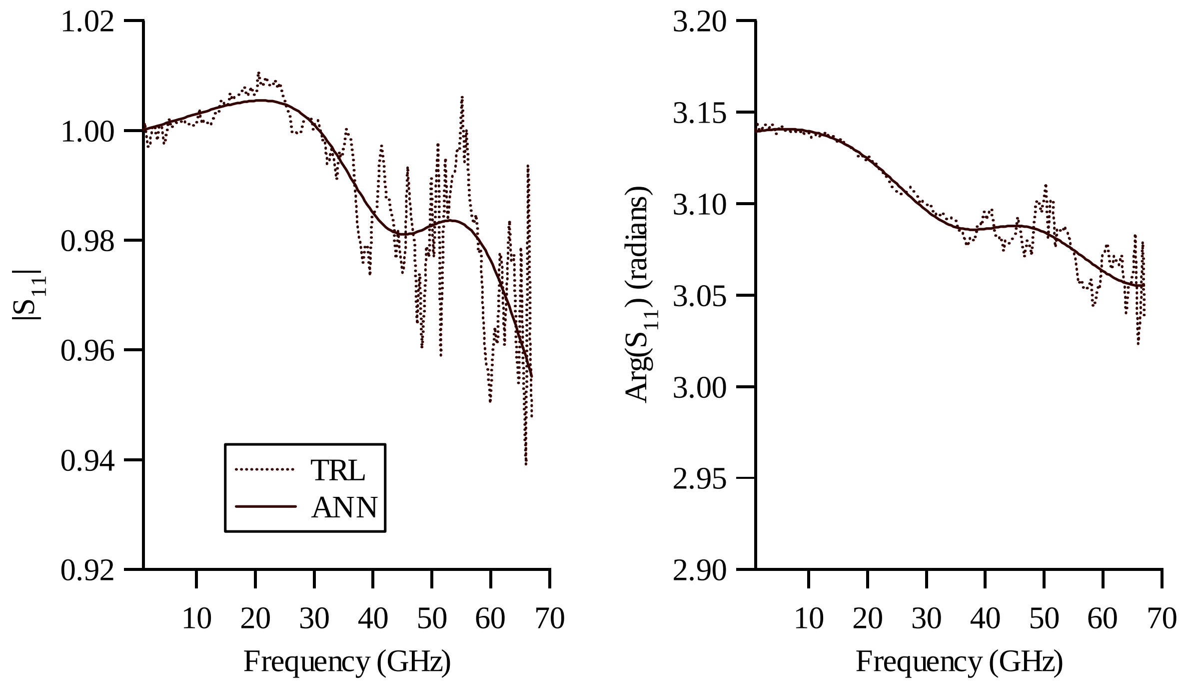

Figure 3. Magnitude and phase of S_{11} for the short standard measured by multiline TRL and modeled by an ANN.

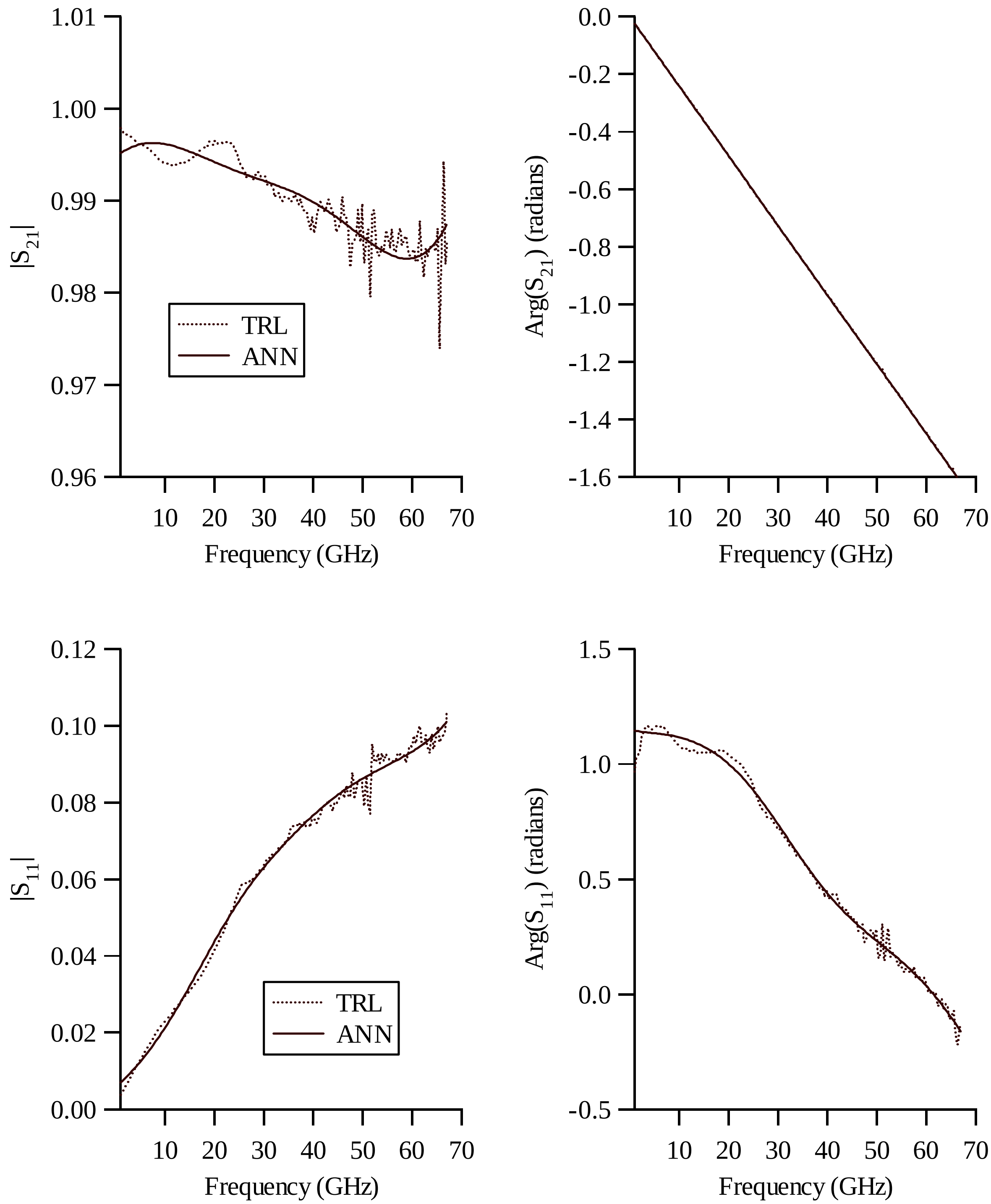

Figure 4. Magnitude and phase of S_{21} and S_{11} for the thru standard measured by multiline TRL and modeled by ANNs.

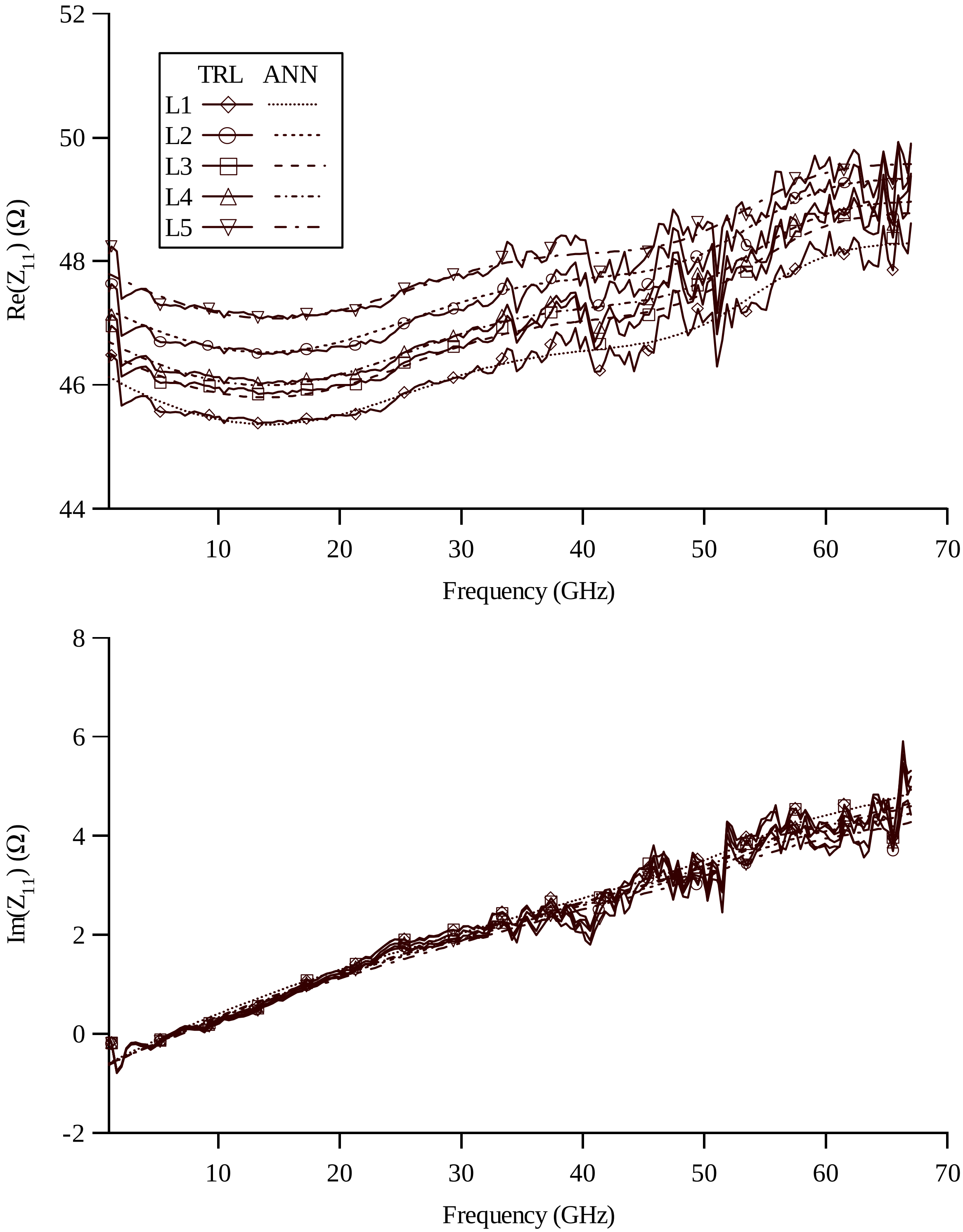

Figure 5. Real and imaginary components of Z_{11} for the load standards measured by multiline TRL and modeled by an ANN.

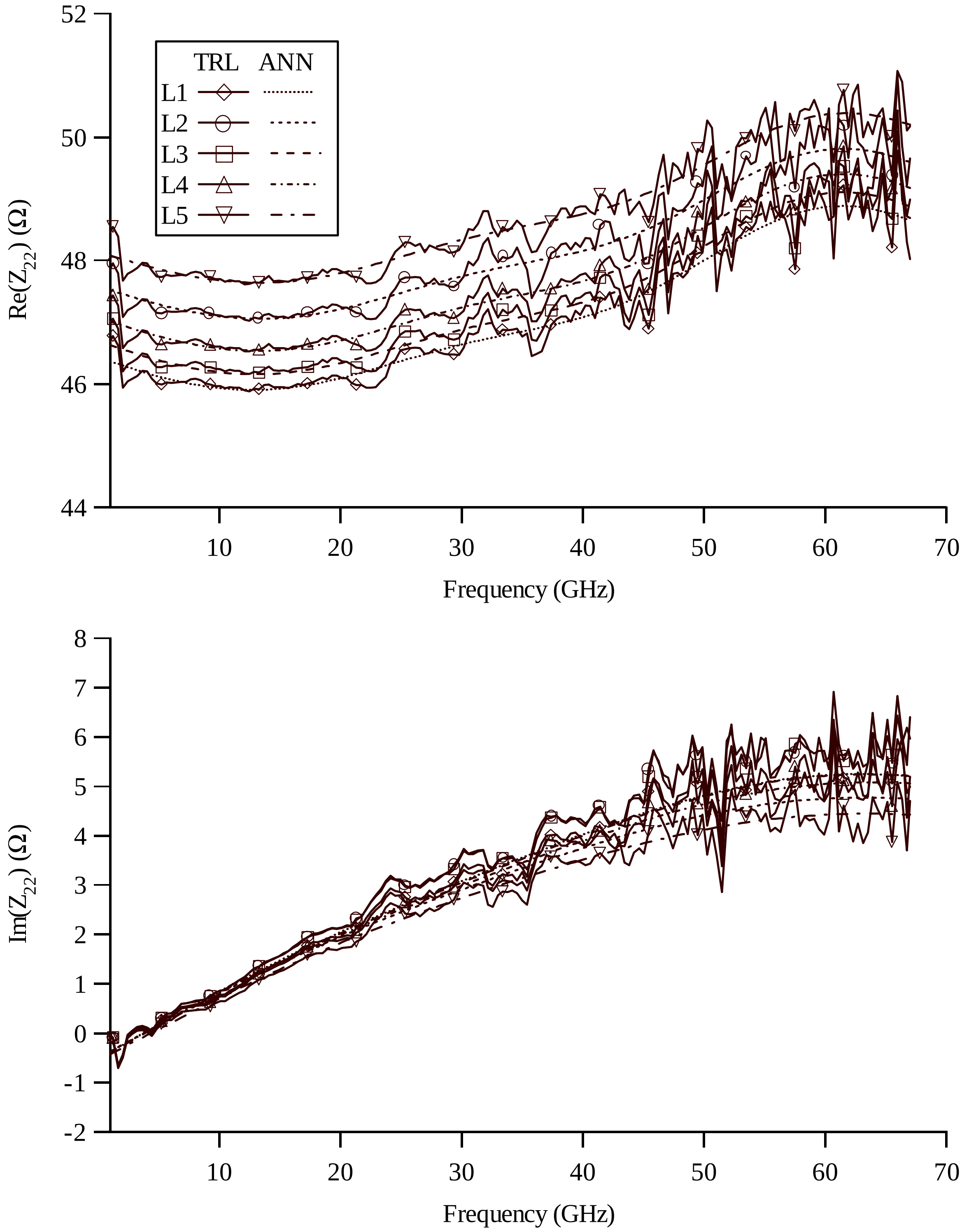

Figure 6. Real and imaginary components of Z_{22} for the load standards measured by multiline TRL and modeled by an ANN.

Table 1. Measured DC resistances of the five load terminations.

Load	DC Resistance (Ω) Port 1	DC Resistance (Ω) Port 2
1	44.73	45.01
2	45.85	46.13
3	45.20	45.27
4	45.38	45.64
5	46.45	46.71

IV. ADVANTAGES OF ANN MODELS

One of the advantages of using ANN models as opposed to calibrated measurement files is the compact description possible with an ANN. For example, the ANN model we developed for the load at port 1 required 62 real-valued parameters to generate complex s-parameters as a function of frequency and DC resistance. In contrast, a single measurement file contains 495 real-valued numbers (165 frequency points plus the real and imaginary components at each point). If a measurement database of just 5 loads is utilized, the combined files would contain 2475 real-valued numbers.

In addition to the size advantage of ANN models, they are also easy to train and use. Detailed physical descriptions and equivalent circuit models are avoided. Once the OSLT standards are measured, one of a number of commercially available ANN programs may be used to model the standards. We used software that is available, free of charge, from Zhang et al. [16] to construct our ANN models. After the model is trained, it can be exported as line code and used in custom software that performs OSLT calibrations [2,17].

We explored the accuracy of ANN models trained with only a few measurement points. We did this by training an ANN model at port 1 using the same 3 loads (1, 4, and 5) but this time we used only 9 of the 165 measurement points. By purposely not training the ANN model at all the available frequencies, we could test how effective the model behaved at the other 156 frequencies. We found that the ANN model trained at only 9 points exhibited almost identical deviations between measured and predicted values as the ANN model trained at all 165 points. Our observation that so few training points are sufficient to model our standards highlights another important advantage in using ANN models over calibrated measurement data files. We found that it is possible to cut down on calibration times by measuring only a few frequency points and developing an ANN model, rather than measuring numerous points and storing large data files.

We also explored the use of ANN models for extrapolation outside the bounds of the training data. (Generally, it is believed that ANN models are good at interpolating but not extrapolating.) We did this by training an ANN model at port 1 using 3 of the 5 loads once again, but this time we chose loads 2, 3, and 4. By purposely not training the ANN model with loads 1 and 5, we could test how effective the model behaved at extrapolating. Surprisingly, both the interpolating and extrapolating ANN models exhibited almost identical deviations between measured and predicted values. This bodes well for the application of ANN models to our loads, since it is conceivable that other wafers may possess DC resistances slightly outside the range of the 5 loads we used to train the models.

V. CALIBRATION COMPARISONS

We performed 5 OSLT calibrations, each one making use of the same ANN-modeled open, short, and thru standards as well as the ANN-modeled loads with their respective DC resistances. We calibrated a 1.764-mm long CPW transmission line using each of the OSLT calibrations and compared the results to measurements calibrated directly using the benchmark multiline TRL calibration. Figure 7 compares the magnitudes of S_{21} and S_{11} for all 6 calibrations. The agreement is remarkably good except at a few points where the multiline TRL calibration is extremely noisy.

To obtain a more quantitative idea of the differences, we plotted the maximum magnitude of the vector differences of the scattering parameters [max($|S_{ij}|$)] for the 1.764-mm line for each of the OSLT calibrations and the multiline TRL calibration. Figure 8 illustrates the differences. All of the OSLT calibrations using ANN-modeled standards compare favorably to the benchmark multiline TRL calibration, with a difference of less than 0.04 in magnitude at most of the frequencies over the 66 GHz bandwidth. Not surprisingly, the OSLT calibrations for loads 2 and 3 show slightly higher differences since they were not used to train the ANN model. The differences between the 5 OSLT calibrations and the TRL calibration do not necessarily mean the OSLT calibrations are in error. The differences are likely due to the presence of noise in the TRL calibration that the ANN models avoided. Regardless of the source of error, a 0.04 difference between two on-wafer calibrations spanning 66 GHz is impressive, considering that the repeatability between two multiline TRL calibrations is usually on the same order.

VI. CONCLUSIONS

In practice, ANN-modeled calibration standards can be easily implemented using existing or custom software packages. In our case, we utilized MultiCal, a freely-available program developed by the National Institute of Standards and Technology, to perform our benchmark multiline TRL calibration. The internal software on any commercial network analyzer can also be used, if the user has confidence in another calibration method such as single-line TRL or LRM (line-reflect-match). Then, once the OSLT standards are measured, one of a number of ANN programs may be used to model the standards. We used software developed by Zhang et al. [16] to construct our ANN models. Finally, a program that can perform OSLT calibrations using exported ANN models is required. We wrote custom software to perform this task, using the equations found in references [2] and [17] to perform the OSLT calibrations.

We have successfully applied ANNs to model the correlation between DC resistance and RF variations in load terminations and the RF performance of open, short,

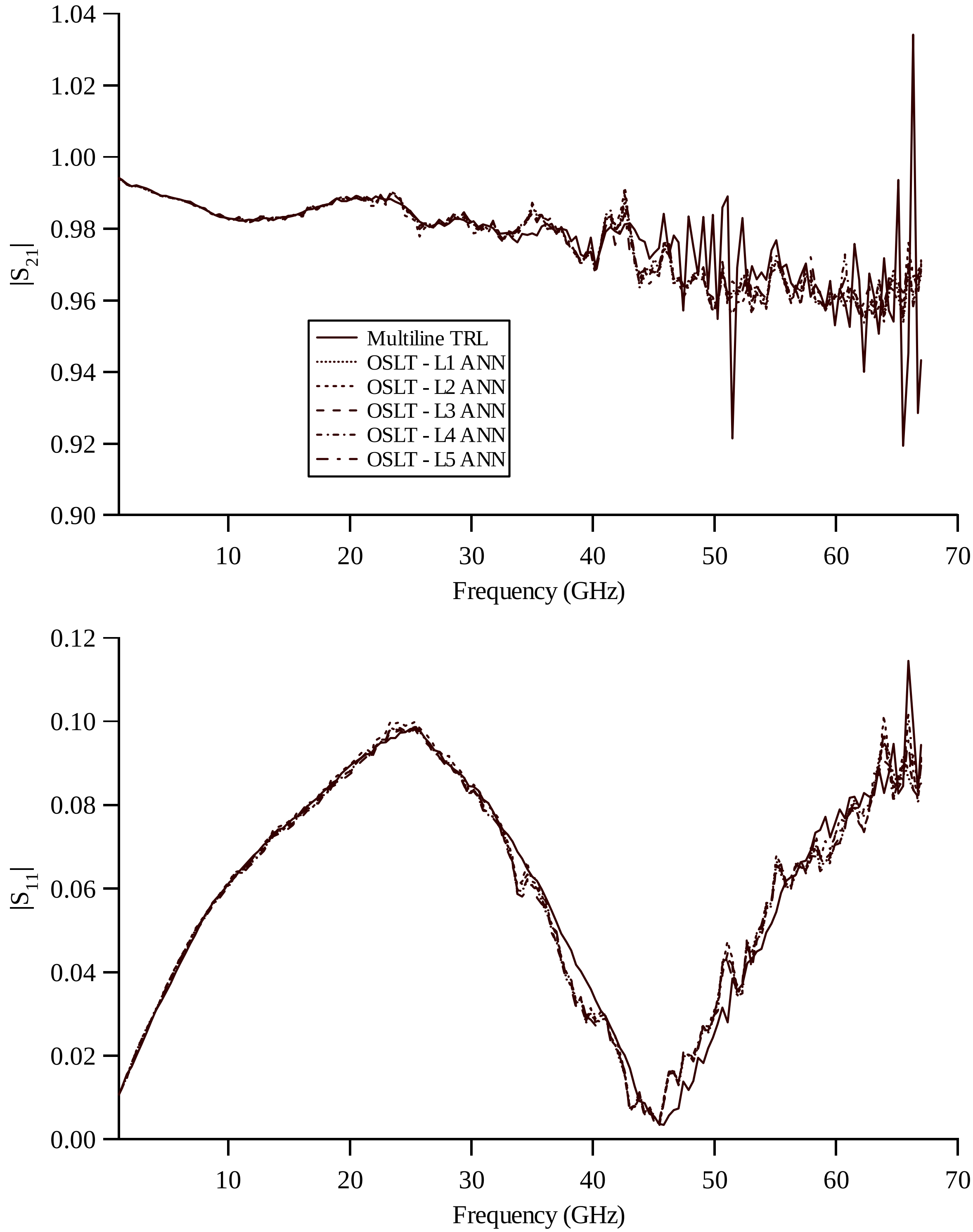

Figure 7. Magnitudes of S_{21} and S_{11} for a calibrated 1.764 mm CPW transmission line.

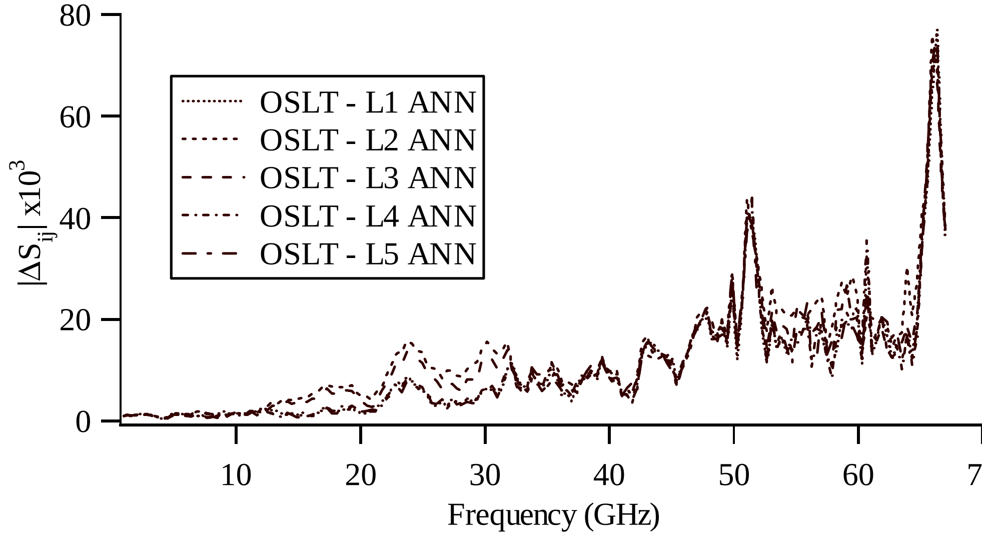

Figure 8. Magnitude of the scattering parameter differences of a calibrated 1.764-mm CPW transmission line.

and thru standards used for on-wafer OSLT calibrations of vector network analyzers. We have shown that these modeled standards compare favorably (a difference of less than 0.04 in magnitude at most frequencies) to the benchmark multiline TRL calibration over a 66 GHz bandwidth.

We have shown that ANN models offer a number of advantages over the use of calibrated measurement files or equivalent circuit models, namely, the following: (1) they do not require detailed physical descriptions, (2) calibration times can be reduced since only a few training points are required to accurately model the standards, (3) ANN model descriptions are much more compact than large measurement data files, (4) they eliminate noise inherent in measured data, and (5) ANN models are able to accurately model loads with measured DC resistances slightly outside their training range.

REFERENCES

[1] R. B. Marks, A multiline method of network analyzer calibration, IEEE Trans Microwave Theory Tech 39 (1991), pp. 1205-1215.

[2] D. K. Rytting, Network analyzer error models and calibration methods, 52nd ARFTG Conference, Short Course on Computer-Aided RF and Microwave Testing and Design, Rohnert Park, CA, Dec. 1998.

[3] J. A. Jargon, K. C. Gupta, and D. C. DeGroot, Artificial neural network modeling for improved on-wafer OSLT calibration standards, Int J RF and Microwave CAE 10 (2000).

[4] P. Kirby, L. Dunleavy, and T. Weller, The effect of load variations on on-wafer lumped element based calibrations, 54th ARFTG Conf Dig, Atlanta, GA, Dec. 1999, pp. 81-90.

[5] K. C. Gupta, EM-ANN models for microwave and millimeter-wave components, IEEE Int Microwave Symp, Workshop on Applications of ANNs to Microwave Design, Jun. 1997, pp. 17-47.

[6] Q. J. Zhang and K. C. Gupta, Neural networks for RF and microwave design, Artech House, Boston, London, 2000.

[7] A. Patnaik and R. K. Mishra, ANN techniques in microwave engineering, IEEE Microwave Magazine 1 (2000), pp. 55-60.

[8] F. Wang, V. K. Devabhaktuni, C. Xi, and Q. J. Zhang, Neural network structures and training algorithms for RF and microwave applications, Int J RF and Microwave CAE 9 (1999), pp. 216-240.

[9] P. M. Watson and K. C. Gupta, EM-ANN models for microstrip vias and interconnects in dataset circuits, IEEE Trans Microwave Theory Tech 44 (1996), pp. 2495-2503.

[10] A. Veluswami, M. S. Nakhla, and Q. J. Zhang, The application of neural networks to EM-based simulation and optimization of interconnects in high-speed VLSI circuits, IEEE Trans Microwave Theory Tech 45 (1997), pp. 712-723.

[11] G. L. Creech, B. J. Paul, C. D. Lesniak, T. J. Jenkins, and M. C. Calcatera, Artificial neural networks for fast and accurate EM-CAD of microwave circuits, IEEE Trans Microwave Theory Tech 45 (1997), pp. 794-802.

[12] F. Wang and Q. J. Zhang, Knowledge-based neural models for microwave design, IEEE Trans Microwave Theory Tech 45 (1997), pp. 2333-2343.

[13] P. M. Watson and K. C. Gupta, Design and optimization of CPW circuits using EM-ANN models for CPW components, IEEE Trans Microwave Theory Tech 45 (1997), pp. 2515-2523.

[14] A. Patnaik, R. K. Mishra, G. K. Patra, and S. K. Dash, An artificial neural network model for effective dielectric constant of microstrip line, IEEE Trans Antennas Propagation 45 (1997), p. 1697.

[15] M. Vai, S. Wu, B. Li, and S. Prasad, Creating neural network based microwave circuit models for analysis and synthesis, Proc Asia Pacific Microwave Conf, Hong Kong, Dec. 1997, pp. 853-856.

[16] NeuroModeler, ver. 1.2, Q. J. Zhang and his neural network research team, Department of Electronics, Carleton University, Ottawa, Canada, 1999.

[17] D. C. DeGroot, K. L. Reed, and J. A. Jargon, Equivalent circuit models for coaxial OSLT standards, 54th ARFTG Conference Digest, Atlanta, GA, Dec. 1999, pp. 103-115.